

WHAT IS CLAIMED IS:

1. A method for compressing data, comprising the steps of:
analyzing a data block of an input data stream to identify a data type of the data block,
5 the input data stream comprising a plurality of disparate data types;
performing content dependent data compression on the data block, if the data type of the
data block is identified;
performing content independent data compression on the data block, if the data type of
the data block is not identified.
10
2. The method of claim 1, wherein the step of performing content independent data
compression, comprises:
encoding the data block with a plurality of encoders to provide a plurality of encoded
data blocks;
15 determining a compression ratio obtained for each of the encoders;
comparing each of the determined compression ratios with a first compression threshold;
selecting for output the input data block and appending a null compression descriptor to
the input data block, if all of the encoder compression ratios do not meet the first compression
threshold; and
20 selecting for output the encoded data block having the highest compression ratio
and appending a corresponding compression type descriptor to the selected encoded data block,
if at least one of the compression ratios meet the first compression threshold.
3. The method of claim 1, wherein the step of performing content dependent
25 compression comprises the steps of:
selecting one or more encoders associated with the identified data type and encoding the
data block with the selected encoders to provide a plurality of encoded data blocks;
determining a compression ratio obtained for each of the selected encoders;
comparing each of the determined compression ratios with a second compression
30 threshold;

selecting for output the input data block and appending a null compression descriptor to the input data block, if all of the encoder compression do not meet the second compression threshold; and

selecting for output the encoded data block having the highest compression ratio
5 and appending a corresponding compression type descriptor to the selected encoded data block, if at least one of the compression ratios meet the second compression threshold.

4. The method of claim 1, wherein the step of performing content independent data compression on the data block, if the data type of the data block is not identified, comprises the
10 steps of:

estimating a desirability of using of one or more encoder types based one characteristics of the data block; and

compressing the data block using one or more desirable encoders.

15 5. The method of claim 1, wherein the methods steps are tangibly embodied as program instructions on a program storage device, wherein the instructions are executable by a machine to perform the method steps.

6. The method of claim 1, wherein the step of performing content dependent data
20 compression on the data block, if the data type of the data block is identified, comprises the steps of:

estimating a desirability of using of one or more encoder types based on characteristics of the data block; and

compressing the data block using one or more desirable encoders.

25

7. The method of claim 1, wherein the step of analyzing the data block comprises one of analyzing the data block to recognize one of a data type, data structure, data block format, file substructure, file types and a combination thereof.

30

8. The method of claim 7, further comprising the step of maintaining an association between encoder types and one of data types, data structures, data block formats, file substructure, file types and a combination thereof.

5 9. A method for compressing data, comprising the steps of:
analyzing a data block of an input data stream to identify a data type of the data block,
the input data stream comprising a plurality of disparate data types;
performing content dependent data compression on the data block, if the data type of the
data block is identified;
10 determining a compression ratio of the compressed data block obtained using the content
dependent compression and comparing the compression ratio with a first compression threshold;
and
performing content independent data compression on the data block, if the data type of
the data block is not identified or if the compression ratio of the compressed data block obtained
15 using the content dependent compression does not meet the first compression threshold.

10. The method of claim 9, wherein the step of performing content independent data
compression, comprises:
encoding the data block with a plurality of encoders to provide a plurality of encoded
20 data blocks;
determining a compression ratio obtained for each of the encoders;
comparing each of the determined compression ratios with a second compression
threshold;
selecting for output the input data block and appending a null compression descriptor to
25 the input data block, if all of the encoder compression ratios do not meet the second compression
threshold; and
selecting for output the encoded data block having the highest compression ratio
and appending a corresponding compression type descriptor to the selected encoded data block,
if at least one of the compression ratios meet the second compression threshold.

30

11. The method of claim 9, wherein the step of performing content dependent compression comprises the steps of:

selecting one or more encoders associated with the identified data type and encoding the data block with the selected encoders to provide a plurality of encoded data blocks;

5 determining a compression ratio obtained for each of the selected encoders;

comparing each of the determined compression ratios with the first compression threshold;

selecting for output the input data block and appending a null compression descriptor to the input data block, if all of the encoder compression do not meet the first compression
10 threshold; and

selecting for output the encoded data block having the highest compression ratio and appending a corresponding compression type descriptor to the selected encoded data block, if at least one of the compression ratios meet the first compression threshold.

12. The method of claim 9, wherein the step of performing content dependent data compression when the data type of the data block is identified, further comprises the steps of:

estimating a desirability of using of one or more encoder types based one characteristics of the data block; and

compressing the data block using one or more desirable encoders.

20

13. The method of claim 9, wherein the step of performing content independent data compression on the data block, if the data type of the data block is not identified, comprises the steps of:

estimating a desirability of using of one or more encoder types based one characteristics
25 of the data block; and

compressing the data block using one or more desirable encoders.

14. The method of claim 9, wherein the methods steps are tangibly embodied as program instructions on a program storage device, wherein the instructions are executable by a
30 machine to perform the method steps.

15. The method of claim 9, wherein the step of analyzing the data block comprises one of analyzing the data block to recognize one of a data type, data structure, data block format, file substructure, file types and a combination thereof.

5

16. The method of claim 15, further comprising the step of maintaining an association between encoder types and one of data types, data structures, data block formats, file substructure, file types and a combination thereof.

10 17. A program storage device readable by a machine, tangibly embodying a program instructions executable by the machine to perform method steps for compressing data, the method steps comprising:

receiving an input data stream comprising a plurality of disparate data types;

compressing the input data stream using each of a plurality of different encoders; and

15 generating an encoded data stream by selectively combining compressed data blocks output from each of the encoders based on compression ratios obtained by the encoders.

18. The program storage device of claim 17, wherein the instructions for performing the step of generating the encoded data stream comprise instructions for tagging each compressed
20 data block with a compression type descriptor.

19. The program storage device of claim 17, wherein the step of generating the encoded data stream comprises combining uncompressed data blocks from the input data stream with the compressed data blocks and tagging each uncompressed data block with a null compression
25 descriptor.

20. The program storage device of claim 17, wherein the instructions for performing the step of compressing the input data stream comprise instructions for compressing each data block in the input data stream using each of the encoders, and wherein the instructions for
30 performing the step of generating the encoded data stream comprise instructions for:

for each data block in the input stream,

determining a compression ratio obtained from each of the encoders;

selecting for output the input data block and appending a null compression descriptor to input the data block, if no compression ratio exceeds a predetermined threshold; and

selecting for output the encoded data block having the greatest compression ratio associated therewith that meets the predetermined threshold and appending a compression type descriptor to the selected encoded data block.

21. The program storage device of claim 20, further comprising instructions for performing the step of applying a predetermined timing constraint to the compression process to provide real-time data compression of the input data stream.

22. The program storage device of 21, wherein the instructions for performing the step of applying a predetermined time constraint comprise instructions for performing the steps of: initializing a timer with a user-specified time interval upon commencing compression of an input data block; and

terminating the encoding step upon the earlier of one of the expiration of the timer and the completion of the encoding of the input data block by all of the plurality of encoders;

wherein the step of determining the compression ratios is only performed for the encoders that have completed encoding of the input data block before expiration of the timer.